DOCUMENT RESUME

ED 173 368 TM 008 559

AUTHOR Serlin, Ronald C.; Marascuilo, Leonard A.
TITLE Omnibus Tests for Interactions in Repeated Measures Designs with Dichotomous Dependent Variables.
INSTITUTION California Univ., Berkeley. School of Education.
PUB DATE [Mar 78]
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Hypothesis Testing; Mathematical Models; *Nonparametric Statistics; Research Design; Simulation; *Statistical Analysis; *Tests of Significance
IDENTIFIERS *Omnibus Tests; *Repeated Measures Designs

ABSTRACT

When examining a repeated measures design with independent groups for a significant group by trial interaction, classical analysis of variance or multivariate procedures can be used if the assumptions underlying the tests are met. Neither procedure may be justified for designs with small sample sizes and dichotomous dependent variables. An omnibus nonparametric test statistic for interactions in such designs, based on the alignment procedures of Hodges and Lehmann, is presented and examined. The results indicate that this statistic should be preferred to the F test when sample sizes are small and the probabilities are near 0.0 or 0.1. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED173368

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Omnibus Tests for Interactions in Repeated Measures Designs with Dichotomous Dependent Variables[1]

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Ronald C. Serlin

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Ronald C. Serlin

Leonard A. Marascuilo

University of California, Berkeley

[1]This research was supported by funds made available to the School of Education by the Computer Center, Campbell Hall, University of California, Berkeley.

TM008 559

Abstract.

When examining a repeated measures design with independent groups for a significant group by trial interaction, classical ANOVA or multivariate procedures can be used if the assumptions underlying the tests are met. Neither procedure may be justified for designs with small sample sizes and dichotomous dependent variables. An omnibus nonparametric test statistic for interactions in such designs, based on the alignment procedures of Hodges and Lehmann (1962), is presented and examined. The results indicate that this statistic should be preferred to the F test when sample sizes are small and the probabilities are near 0.0 or 1.0.

Introduction: Interactions and Dichotomous Dependent Variables.

Research questions considered by behavioral scientists often center on the evaluation of interactions between two or more independent variables. In the completely independent group model based on normality and common variace assumptions, standard ANOVA procedures can be called upon to study interaction effects in an efficient manner. Statistically significant interactions can be examined in detail in terms of linear contrasts either in a post hoc or an a priori way according to the methods described by Marascuilo and Levin (1970).

Interactions for repeated measures also can be examined by planned contrasts or omnibus F tests with standard ANOVA procedures if errors are normally distributed with equal variances and covariances. If these conditions are not satisfied, multivariate analysis of variance models can be used. Neither of these models may be justified for small samples and dichotomous dependent variables, since small sample data cannot be tested for possible assumption violations and since both procedures require a quantitative variable with, ideally, a continuous probability distribution.

Lunney (1970) investigated the two-factor F test of no interaction for a dichotomous dependent variable observed on independent groups of subjects and found that, if the total number of observations is small, the actual type I error rate is smaller than nominal. Without showing results, he also reports that the power for dichotomous variables in ANOVA is much smaller than the theoretical values for small samples based on normal curve assumptions. McSweeney (1968) examined the same model as Lunney and

proposed a nonparametric test for interaction in which ranks replace the original observations after the data are aligned, according to the methods described by Hodges and Lehmann (1962). She found that her version of the alignment interaction test generated a liberal type I error rate for small sample sizes and for small values of the noncentrality parameter. In addition, power comparisons for small samples favored her alignment test; however, this advantage was lost on large samples.

Neither of these investigators examined the problem of measuring interactions in repeated measures designs with a dichotomous dependent variable. Because repeated measures involve a larger number of observations per subject or block, one might expect more favorable type I and type II error rates for dichotomous variables, even for small sample sizes within groups.

If the dependent variable is dichotomous and if there are only two observation periods, the general repeated measures design is similar to McNemar's test of change (1947) involving multiple groups, rather than one group. This multiple group design has been modified by Marascuilo (1976) to provide a test of two factor interactions. Hamdan, Pirie, and Arnold (1975) have also suggested tests of the hypothesis of no interaction or the hypothesis of equal change parameters in the multiple sample McNemar data model, as have Bishop, Fienberg, and Holland (1975). For more than two repeated measures per subject, Marascuilo and Serlin (1977) have provided a method using planned contrasts for examining interactions in a multiple sample Cochran test model. As yet, no single omnibus test statistic has been suggested and examined which analyzes the group by trial interaction in a re-

peated measures design with dichotomous dependent variables. In this article, a nonparametric omnibus test statistic based upon the alignment procedures of Hodges and Lehmann is proposed, and properties of this statistic are examined.

Rationale for the Aligned Test of No Interaction.

The basic multiple sample repeated measures design under consideration can be represented as follows. For the orthogonal case with equal cell sizes, N, let X_{ijk} denote the observation (here equal to 0 or 1) for the i^{th} subject in the j^{th} group, G_j, on the k^{th} trial, T_k, where $j \leq J$, $k \leq K$. The latent structure model for this design is given by:

$$X_{ijk} = \mu + \alpha_{ij.} + \varepsilon_{jk(i)}$$

where:

$$\varepsilon_{jk(i)} = \alpha_{..k} + \gamma_{.jk} + \varepsilon_{ijk}$$

and:

$$\alpha_{ij.} = \alpha_{.j.} + \ell_{ij.}$$

so that;

$$X_{ijk} = \mu + \alpha_{.j.} + \ell_{ij.} + \alpha_{..k} + \gamma_{.jk} + \varepsilon_{ijk}$$

It is worth noting that this structure applies to both the multiple sample McNemar and Cochran Q test data models described in the previous section. Indeed, Lehmann (1967) has shown that the Cochran Q test model (1950) reduces to the McNemar test for one group and two repeated measurements.

The standard repeated measures F test of no interactions based on normality assumptions is given by $F = MS_{G\times T}/MS_{\varepsilon}$. Even if the assumptions are not satisfied, one can perform this analysis on the raw dichotomous data. The only probelm is that

the test statistic may not follow the theoretical F and non-central F distributions. An alternative test based on ranks that might have better distribution properties can be derived from the alignment procedures of Hodges and Lehmann. For this test the latent structure equation can be rearranged to yield "aligned" scores whose expected values are estimates of the interaction effects. To achieve the alignment, one removes the group, trial, and subject effects by adjusting the observations on the sample mean of the j^{th} group, the k^{th} observation period, and the i^{th} subject to obtain:

$$\hat{X}^{A}_{ijk} = X_{ijk} - \hat{\alpha}_{.j.} - \hat{\alpha}_{..k} - \hat{\ell}_{ij.} = \hat{\mu} + \hat{\gamma}_{.jk} + \hat{\varepsilon}_{ijk}$$

This alignment can be achieved by subtracting the k^{th} trial mean and the i^{th} subject mean from X_{ijk}. The aligned observations are then replaced by their ranks from 1 to NJK (smallest to largest) and the analysis proceeds on the basis of these ranks, R_{ijk}. Since the raw data are dichotomous, there will probably be many tied aligned observations. Average ranks are used for the tied values. Finally, the test statistic is computed as:

$$X^2 = \frac{J(N-1)(K-1)SS^{(R)}_{\text{groups} \times \text{trials}}}{SS^{(R)}_{\text{subjects} \times \text{group} \times \text{trials}}}$$

where the sum of squares is computed in terms of the ranks and not the original observed values. Thus:

$$SS^{(R)}_{\text{group} \times \text{trials}} = \frac{1}{N}\sum_j\sum_k\left(\sum_i R_{ijk}\right)^2 - \frac{1}{NK}\sum_j\left(\sum_k\sum_i R_{ijk}\right)^2$$

$$- \frac{1}{NJ}\sum_k\left(\sum_j\sum_i R_{ijk}\right)^2 + \frac{1}{NJK}\left(\sum_j\sum_k\sum_i R_{ijk}\right)^2$$

$$SS^{(R)}_{\text{subject} \times \text{group} \times \text{trials}} = \sum_j \sum_k \sum_i R^2_{ijk} - \frac{1}{K} \sum_j \sum_i \left(\sum_k R_{ijk} \right)^2$$

$$- \frac{1}{N} \sum_j \sum_k \left(\sum_i R_{ijk} \right)^2 + \frac{1}{NK} \sum_j \left(\sum_k \sum_i R_{ijk} \right)^2$$

With this formulation, X^2 has an asymptotic Chi-square distribution with $\nu = (J - 1)(K - 1)$ degrees of freedom. For the F test based on the raw data, the test for interaction has $\nu_1 = (J - 1)(K - 1)$ and $\nu_2 = J(N - 1)(K - 1)$ degrees of freedom. The closeness of the distribution of X^2 to χ^2 for small samples is shown in the following sections.

Generation of Samples and Method of Analysis.

Since the dependent variable is dichotomous, for K trials there are only 2^K combinations of 0's and 1's possible for each subject's outcome. At the same time, the score for a subject on trial k is related to his score on trial k'. A between trial correlation coefficient, ρ, equal to 0.50 was chosen to ensure that the data generated were substantially different from an independent observation model. This value for ρ was used throughout.

Table 1 indicates the eight response combinations possible for three repeated measures. If p_{jk} represents the proba-

-----Insert Table 1 about here-----

bility of a subject in the j^{th} group achieving a score of 1 on the k^{th} trial, then the relative proportions of each of the eight response combinations can be derived such that $E(X_{ijk}) = p_{jk}$ and that the expected between trial correlation coefficient equals 0.5. Let the proper relative proportion of the r^{th} response combination be denoted by a_r, where $1 \leq r \leq 2^K$. Then, in order to achieve the correct values for p_{jk} and ρ, we must have:

$$a_1 + a_2 + a_3 + a_4 = p_{j1}$$

$$a_1 + a_2 + a_5 + a_6 = p_{j2}$$

$$a_1 + a_3 + a_5 + a_7 = p_{j3}$$

and:

$$\frac{a_1 + a_2 - p_{j1}p_{j2}}{\sqrt{p_{j1}(1 - p_{j1})p_{j2}(1 - p_{j2})}} = \rho$$

$$\frac{a_1 + a_3 - p_{j1}p_{j3}}{\sqrt{p_{j1}(1 - p_{j1})p_{j3}(1 - p_{j3})}} = \rho$$

$$\frac{a_1 + a_5 - p_{j2}p_{j3}}{\sqrt{p_{j2}(1 - p_{j2})p_{j3}(1 - p_{j3})}} = \rho$$

As soon as the desired p_{jk} are specified, the a_r can be calculated.

When generating data for three or four trials, the number of equations is not sufficient to specify all weights, a_r. In these cases, values were chosen on a trial-and-error basis until all relevant constraint equations were satisfied. Finally, sample data for each group were selected using these values for a_r and a uniform (0, 1) random number generator. One thousand samples were drawn for a given set of p_{jk} values.

For each set of p_{jk}, samples were generated by a computer and, on each sample, F and X^2 statistics were calculated. By comparing each statistic to the corresponding tabled value at the 0.50, 0.25, 0.10, 0.05, and 0.01 significance levels, the powers of the two tests were determined for a number of different alternatives. The specific null and alternative hypotheses investigated are reported in Table 2. The first 5 cases represent null situa-

-----Insert Table 2 about here-----

tions where there are no interactions and where the probabilities

all equal 0.5. The next five cases represent interactions for which the probabilities are near 0.5. Cases 11 through 15 are null cases for which the probabilities are close to zero or one, and the final six cases involve interactions for such extreme probabilities.

Results.

The results of the 21 Monte Carlo studies are summarized in Tables 3 and 4. Table 3 summarizes the type I error rates

-----Insert Tables 3 and 4 about here-----

for F and X^2 for the 20 null cases generated for N = 5 and N = 10. Table 4 provides the empirical power of these two statistics for the 22 alternative cases based on N = 5 and N = 10.

Type I Error Rates.

As shown, both the F and X^2 statistics are approximately distributed as the corresponding theoretical F and Chi-square distributions, even for N = 5. Small-sample situations for Cases 1 through 5, for which p is near 0.5, are cases that show the proposed test to be somewhat more liberal than F. For the larger-sample situations, the risks of a type I error are not substantially different from the nominal 0.05 value and, in most cases, the fit is adequate at 0.01. However, for the smaller sample size and an extreme probability null case, such as Cases 11, 12, 14, and 15, neither statistic fits the theoretical distribution too closely. Again, for N = 10, the fits for both statistics are much better. While Case 13 is similar to these, it does not show the same type I error rate pattern. In this case, the fit is quite satisfactory, perhaps because the alignment estimators within a subject are based on four trials and, as a result, are more stable.

Power and Type II Error Rates.

Although power values are reported in Table 4 for five critical points, the discussion will focus on powers for type I error rates of 0.05 and 0.01.

Disordinal interactions are represented by the cases whose probabilities are near 0.5. For these, the average ratio of powers, Chi-square to F, was 1.33 for N = 5 and 0.93 for N = 10 at the 0.05 alpha level. For an alpha level of 0.01 and disordinal interactions, the ratios were 3.29 for N = 5 and 1.16 for N = 10. Ordinal interactions are represented by Cases 16 through 21. For these extreme probability situations, the Chi-square test based on alignment shows greater power than the F test at the 0.05 level, with the average ratio of the powers being 1.32 for samples of size 5 and 1.08 for samples of size 10. For an alpha level of 0.01, these ratios were 2.01 for N = 5 and 1.41 for N = 10.

Discussion.

The results of this investigation are generally in agreement with the results of the studies of Lunney and McSweeney. In particular, for small sample sizes and probabilities near 0.0 and 1.0, the risk of a type I error for the proposed aligned test tends to be less than the nominal values, but closer to the nominal values than is the F test. However, this advantage is not evident as the sample size increases to N = 10.

When the probabilities are in the neighborhood of 0.5, the empirical risk of a type I error is closer to the nominal value for the F test than the proposed test, which was somewhat more liberal.

The results of this investigation also show that the

test of interaction for dichotomous variables based on the Hodges-Lehmann alignment procedures tends to have more power than the corresponding F test based on the original (0, 1) variables. This tendency is most pronounced when the probabilities are close to 0.0 or 1.0.

The greater power is most likely related to the greater number of possible values introduced by ranking the aligned observation from 1 to NJK. As indicated in Tables 3 and 4, the best fits were found for N = 10, J = 2, and K = 4 and for N = 10, J = 4, and K = 2. For these extreme cases, ranks from 1 to 80 were used. Even though there were many tied values, the ranking procedures helped raise the powers. Because of this, one could expect even greater increases in power with the use of normal scores and their greater deviations from one another in the tails of the distribution.

In addition, the alignment procedure has the desirable property that, for one group and two trials, the associated test statistic is asymptotically equivalent to the McNemar test. It can be shown that the alignment χ^2 for this design is given by:

$$\chi_A^2 = \frac{N - 1}{N}\left(\frac{(n_{10} - n_{01})^2}{\left(n_{01} + n_{10} - \frac{(n_{10} - n_{01})^2}{N}\right)}\right)$$

where n_{10} and n_{01} are the frequencies associated with the off-diagonal change cells, and N is the sample size.

Significant interactions can be examined by testing each tetrad difference:

$$\hat{\psi}_{jj'kk'} = \bar{R}_{.jk} - \bar{R}_{.j'k} - \bar{R}_{.jk'} + \bar{R}_{.j'k'}$$

$$j \neq j' \qquad \text{and} \qquad k \neq k'$$

for significance. The hypothesis H_0: $\psi_{jj'kk'} = 0$ is rejected if:

$$|\hat{\psi}_{jj'kk'}| > S \, SE_{\hat{\psi}}$$

where:

$$S = \sqrt{\chi^2_{(K-1)(J-1):1-\alpha}}$$

and:

$$S^2_{\hat{\psi}} = \frac{4}{N}\left(\frac{SS^{(R)}_{\text{subjects} \times \text{group} \times \text{trials}}}{J(N-1)(K-1)}\right)$$

If only $Q = \binom{K}{2}\binom{J}{2}$ tetrad differences are of interest, S can be replaced by the N(0, 1) value, $Z_{(1-\alpha/2Q)}$.

On the basis of this investigation, it is reasonable to conclude that the proposed alignment test is to be preferred f small samples whenever the probabilities are near 0.0 or 1.0. In these cases, the alignment test's type I error rates are close to nominal and the power greater than the F test. On the other hand, it seems to make little difference which test is used when the probabilities fluctuate around 0.5.

Table 1. Possible Response Combinations for Three Trials.

Combination (r)	Trial One	Trial Two	Trial Three
1	1	1	1
2	1	1	0
3	1	0	1
4	1	0	0
5	0	1	1
6	0	1	0
7	0	0	1
8	0	0	0

Table 2. The 21 Interaction Cases Examined for Closeness of the Fit of F and X^2 to the Theoretical F and χ^2 Distributions.

Probabilities Near 0.5

Case	J	K		T_1	T_2	T_3	T_4
1	2	2	G_1	0.50	0.50		
			G_2	0.50	0.50		
2	2	3	G_1	0.50	0.50	0.50	
			G_2	0.50	0.50	0.50	
3	2	4	G_1	0.50	0.50	0.50	0.50
			G_2	0.50	0.50	0.50	0.50
4	3	2	G_1	0.50	0.50		
			G_2	0.50	0.50		
			G_3	0.50	0.50		
5	4	2	G_1	0.50	0.50		
			G_2	0.50	0.50		
			G_3	0.50	0.50		
			G_4	0.50	0.50		
6	2	2	G_1	0.30	0.70		
			G_2	0.70	0.30		
7	2	3	G_1	0.40	0.50	0.60	
			G_2	0.60	0.50	0.40	
8	2	4	G_1	0.425	0.475	0.525	0.575
			G_2	0.575	0.525	0.475	0.425
9	3	2	G_1	0.40	0.60		
			G_2	0.50	0.50		
			G_3	0.60	0.40		
10	4	2	G_1	0.425	0.575		
			G_2	0.475	0.525		
			G_3	0.525	0.475		
			G_4	0.575	0.425		

Table 2. (Continued)

Probabilities Near 0.0 and 1.0

Case	J	K		T_1	T_2	T_3	T_4
11	2	2	G_1	0.10	0.10		
			G_2	0.90	0.90		
12	2	3	G_1	0.10	0.10	0.10	
			G_2	0.90	0.90	0.90	
13	2	4	G_1	0.10	0.10	0.10	0.10
			G_2	0.90	0.90	0.90	0.90
14	3	2	G_1	0.10	0.90		
			G_2	0.10	0.90		
			G_3	0.10	0.90		
I5	4	2	G_1	0.10	0.90		
			G_2	0.10	0.90		
			G_3	0.10	0.90		
			G_4	0.10	0.90		
16	2	2	G_1	0.10	0.50		
			G_2	0.90	0.50		
17	2	3	G_1	0.10	0.20	0.30	
			G_2	0.90	0.80	0.70	
18	2	4	G_1	0.150	0.200	0.250	0.300
			G_2	0.850	0.800	0.750	0.700
19	3	2	G_1	0.10	0.90		
			G_2	0.20	0.80		
			G_3	0.30	0.70		
20	4	2	G_1	0.175	0.825		
			G_2	0.150	0.850		
			G_3	0.125	0.875		
			G_4	0.100	0.900		
21	4	2	G_1	0.15	0.85		
			G_2	0.20	0.80		
			G_3	0.25	0.75		
			G_4	0.30	0.70		

Table 3. Empirical Type I Error Rates for the Ten Null Cases for Which N = 5 or N = 10.

Probabilities Near 0.5

Case	N		0.500	0.250	0.100	0.050	0.010
1	5	F	0.500	0.268	0.101	0.062	0.013
		X^2	0.510	0.274	0.131	0.082	0.035
	10	F	0.455	0.247	0.099	0.040	0.007
		X^2	0.495	0.257	0.119	0.061	0.016
2	5	F	0.488	0.229	0.093	0.050	0.015
		X^2	0.491	0.252	0.119	0.070	0.031
	10	F	0.483	0.231	0.090	0.047	0.006
		X^2	0.483	0.238	0.104	0.054	0.011
3	5	F	0.451	0.229	0.100	0.052	0.013
		X^2	0.468	0.249	0.128	0.079	0.029
	10	F	0.503	0.240	0.089	0.046	0.011
		X^2	0.508	0.245	0.105	0.059	0.018
4	5	F	0.464	0.218	0.105	0.051	0.013
		X^2	0.499	0.269	0.122	0.085	0.040
	10	F	0.491	0.247	0.099	0.048	0.016
		X^2	0.498	0.252	0.114	0.052	0.020
5	5	F	0.483	0.253	0.111	0.049	0.010
		X^2	0.487	0.270	0.142	0.070	0.023
	10	F	0.485	0.239	0.099	0.057	0.012
		X^2	0.467	0.224	0.102	0.061	0.017

Table 3. (Continued.)

Probabilities Near 0.0 and 1.0

Case	N		0.500	0.250	0.100	0.050	0.010
11	5	F	0.552	0.211	0.040	0.021	0.000
		X^2	0.572	0.213	0.049	0.021	0.008
	10	F	0.487	0.249	0.098	0.039	0.002
		X^2	0.463	0.242	0.063	0.028	0.005
12	5	F	0.577	0.206	0.057	0.020	0.006
		X^2	0.426	0.227	0.097	0.046	0.017
	10	F	0.526	0.255	0.079	0.025	0.006
		X^2	0.444	0.215	0.048	0.017	0.006
13	5	F	0.483	0.210	0.084	0.040	0.002
		X^2	0.508	0.233	0.133	0.059	0.012
	10	F	0.519	0.252	0.103	0.052	0.013
		X^2	0.490	0.251	0.098	0.058	0.017
14	5	F	0.559	0.174	0.030	0.025	0.001
		X^2	0.419	0.174	0.064	0.025	0.021
	10	F	0.480	0.249	0.085	0.053	0.013
		X^2	0.415	0.241	0.086	0.036	0.013
15	5	F	0.462	0.185	0.083	0.021	0.011
		X^2	0.383	0.182	0.062	0.058	0.018
	10	F	0.504	0.251	0.096	0.045	0.013
		X^2	0.470	0.215	0.076	0.045	0.014

Table 4. Empirical Power Probabilities for the 11 Non-Null Cases for Which N = 5 or N = 10.

Probabilities Near 0.5

Case	N		0.050	0.250	0.100	0.050	0.010
6	5	F	0.943	0.864	0.683	0.517	0.150
		X^2	0.948	0.866	0.698	0.524	0.379
	10	F	0.994	0.982	0.930	0.862	0.678
		X^2	0.994	0.973	0.927	0.852	0.703
7	5	F	0.733	0.485	0.237	0.131	0.023
		X^2	0.695	0.516	0.275	0.195	0.079
	10	F	0.831	0.654	0.413	0.297	0.098
		X^2	0.804	0.614	0.412	0.301	0.122
8	5	F	0.643	0.369	0.156	0.073	0.018
		X^2	0.614	0.402	0.201	0.122	0.034
	10	F	0.706	0.457	0.246	0.151	0.050
		X^2	0.680	0.442	0.249	0.163	0.073
9	5	F	0.689	0.420	0.177	0.116	0.013
		X^2	0.644	0.487	0.217	0.141	0.085
	10	F	0.835	0.654	0.451	0.297	0.114
		X^2	0.797	0.591	0.348	0.247	0.118
10	5	F	0.592	0.318	0.161	0.078	0.016
		X^2	0.540	0.267	0.138	0.097	0.033
	10	F	0.736	0.483	0.263	0.157	0.050
		X^2	0.629	0.373	0.188	0.119	0.043

Table 4. (Continued.)

Probabilities Near 0.0 and 1.0

Case	N		0.050	0.250	0.100	0.050	0.010
16	5	F	0.921	0.810	0.606	0.427	0.115
		X^2	0.926	0.812	0.627	0.442	0.292
	10	F	0.992	0.951	0.856	0.753	0.424
		X^2	0.990	0.942	0.851	0.761	0.550
17	5	F	0.872	0.522	0.211	0.107	0.031
		X^2	0.703	0.541	0.252	0.176	0.066
	10	F	0.960	0.854	0.602	0.435	0.135
		X^2	0.907	0.754	0.505	0.348	0.135
18	5	F	0.726	0.387	0.150	0.069	0.019
		X^2	0.572	0.382	0.186	0.095	0.037
	10	F	0.807	0.596	0.358	0.215	0.069
		X^2	0.742	0.516	0.295	0.194	0.072
19	5	F	0.691	0.433	0.221	0.146	0.043
		X^2	0.705	0.443	0.264	0.155	0.076
	10	F	0.746	0.521	0.319	0.202	0.086
		X^2	0.784	0.562	0.353	0.272	0.117
20	5	F	0.569	0.320	0.127	0.075	0.021
		X^2	0.546	0.355	0.169	0.115	0.043
	10	F	0.568	0.325	0.148	0.070	0.010
		X^2	0.574	0.333	0.165	0.093	0.024
21	5	F	0.528	0.304	0.149	0.092	0.029
		X^2	0.506	0.335	0.185	0.119	0.048
	10	F	0.598	0.358	0.176	0.108	0.039
		X^2	0.594	0.355	0.185	0.121	0.052

References.

Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. *Discrete Multivariate Analysis.* Cambridge: MIT Press, 1975.

Cochran, W. The comparison of percentages in matched samples. *Biometrika,* 1950, *37,* 256-266.

Hamdan, M. A., Pirie, W. R., and Arnold, J. C. Simultaneous testing of McNemar's problem for several populations. *Psychometrika,* 1975, *40,* 153-161.

Hodges, J. L., Jr., and Lehmann, E. L. Rank methods for combination of independent experiments in the analysis of variance. *Annals of Mathematical Statistics,* 1962, *33,* 482-497.

Lehmann, E. L. *Nonparametrics: Statistical Methods Based on Ranks.* San Francisco: Holden-Day, Inc., 1975.

Lunney, G. H. Using analysis of variance with a dichotomous dependent variable: An empirical study. *Journal of Educational Measurement,* 1970, *7,* 263-269.

Marascuilo, L. A. Tests and contrasts for comparing change parameters for a multiple sample McNemar data model. Paper presented at the International Congress of Psychology, Paris, 1976.

Marascuilo, L. A. and Levin, J. R. Appropriate post hoc comparisons for interaction and nested hypotheses in analysis of variance designs: The elimination of type IV errors. *American Education Research Journal,* 1970, *7,* 397-421.

Marascuilo, L. A. and Serlin, R. C. Interactions for dichotomous variables in repeated measures designs. *Psychological Bulletin,* 1977, 84, No. 5, 1002-1007.

McNemar, Q. Note on the sampling error of the differences between correlated proportions or percentages. *Psychometrika,* 1947, *12,* 153-157.

McSweeney, M. *An Empirical Study of Two Proposed Nonparametric Tests for Main Effects and Interaction.* Unpublished dissertation thesis, University of California, Berkeley, 1967.